HOW TO ANALYSE POLITICAL DATA?

HOWINDIALIVES.COM

WHAT WE WILL COVER IN THIS PRESENTATION

- What it takes to achieve graphical excellence
- How to analyse political data?
- Sharing our (HowIndiaLives.com) experience in data visualisation
- What does it take to do a good data visualisation/ story?

WHAT I AM

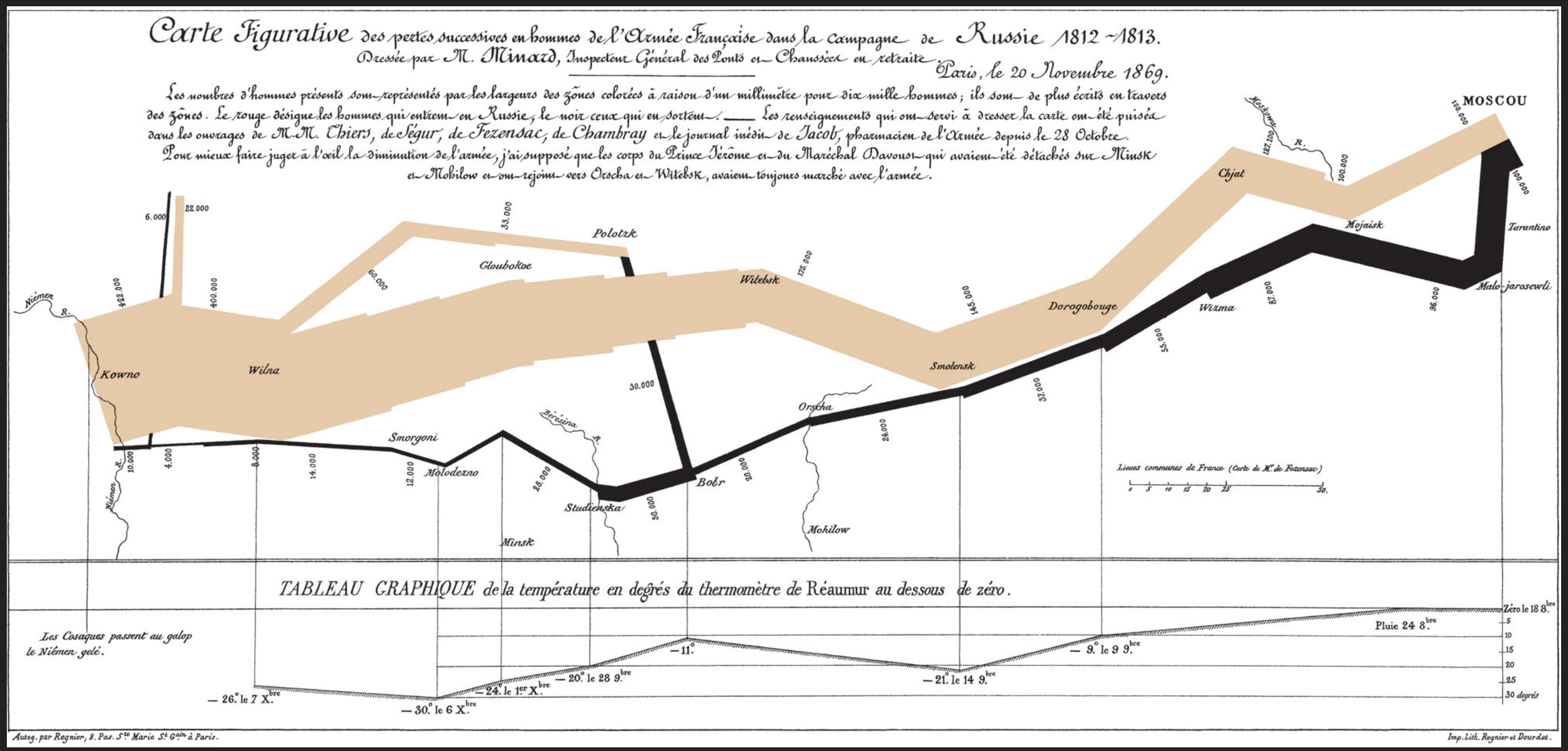
- Journalist with 12 years of experience
- Have used data, especially public data, to write stories
- Co-founded HowIndiaLives.com four years back

WHAT I AM NOT

- Not a software coder
- So this is not a crash course on learning software tools.

SOME CLASSIC EXAMPLES

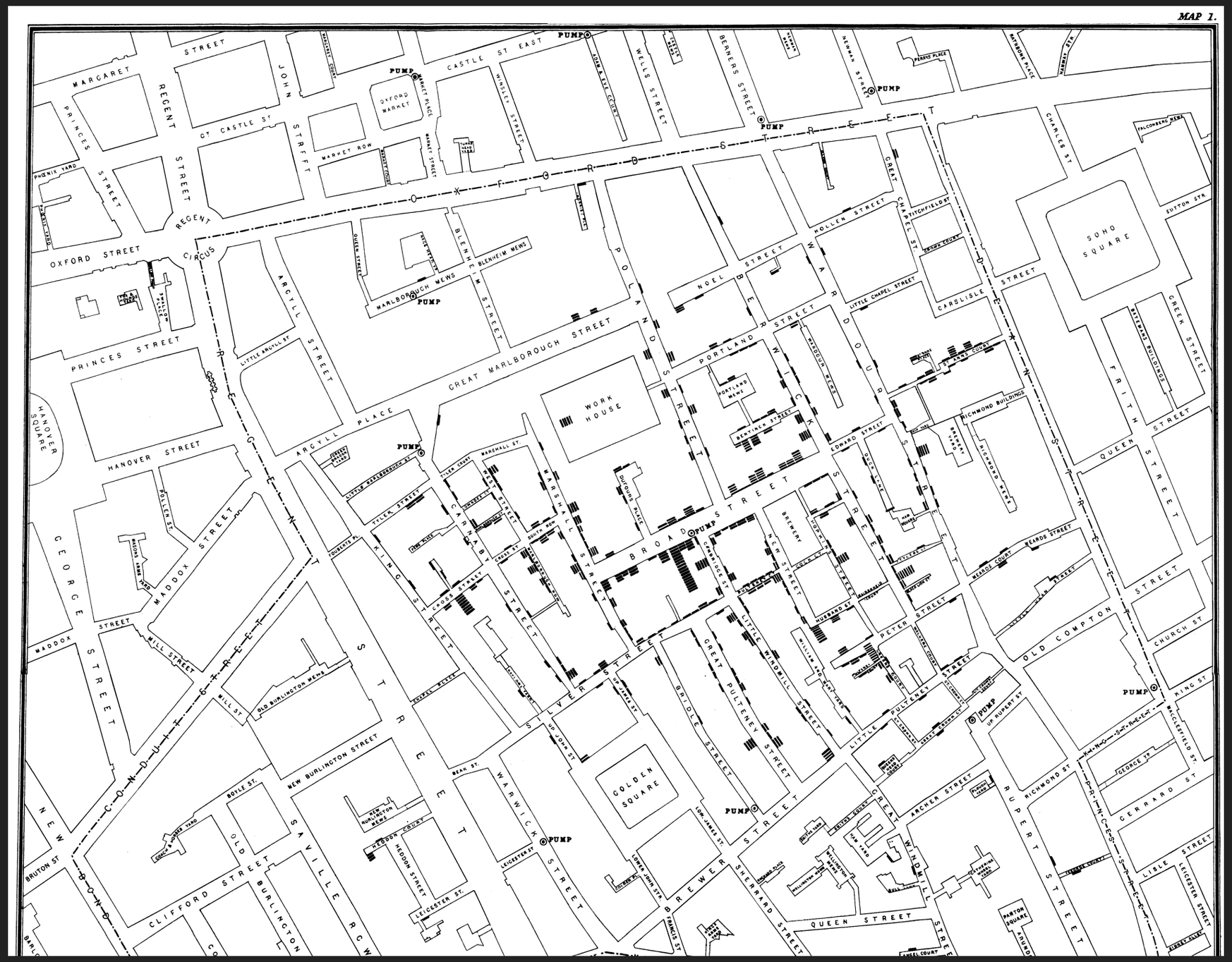
Charles Joseph Minard (1781 - 1870)


Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812 -1813.
Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.
Paris, le 20 Novembre 1869.
Les nombres d'hommes présents sont représentés par les largeurs des zônes colorées à raison d'un millimètre pour dix mille hommes ; ils sont de plus écrits en travers des zônes. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. ___ Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M.M. Thiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.
Pour mieux faire juger à l'oeil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mobilow se sont rejoints vers Orscha et Witebsk, avaient toujours marché avec l'armée.
Niémen R.
Kowno
Wilna
Gloubokoe
Polotzk
Witebsk
Smolensk
Dorogobouge
Chjat
Mojaisk
MOSCOU
Tarantino
Malo-jarosewli
Wixma
Orscha
Berézina R.
Smorgoni
Molodexno
Bobr
Studienska
Minsk
Mohilow
Lieues communes de France (Carte de M.r de Fezensac)
TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.
Les Cosaques passent au galop le Niémen gelé.
— 26.° le 7 X.bre
— 30.° le 6 X.bre
— 24. le 1.er X.bre
— 20.° le 28 9.bre
— 11.°
— 21. le 14 9.bre
— 9.° le 9 9.bre
Pluie 24 8.bre
Zéro le 18 8.bre
Autog. par Regnier, 8. Pas. St.e Marie St Gain à Paris.
Imp. Lith. Regnier et Dourdet.

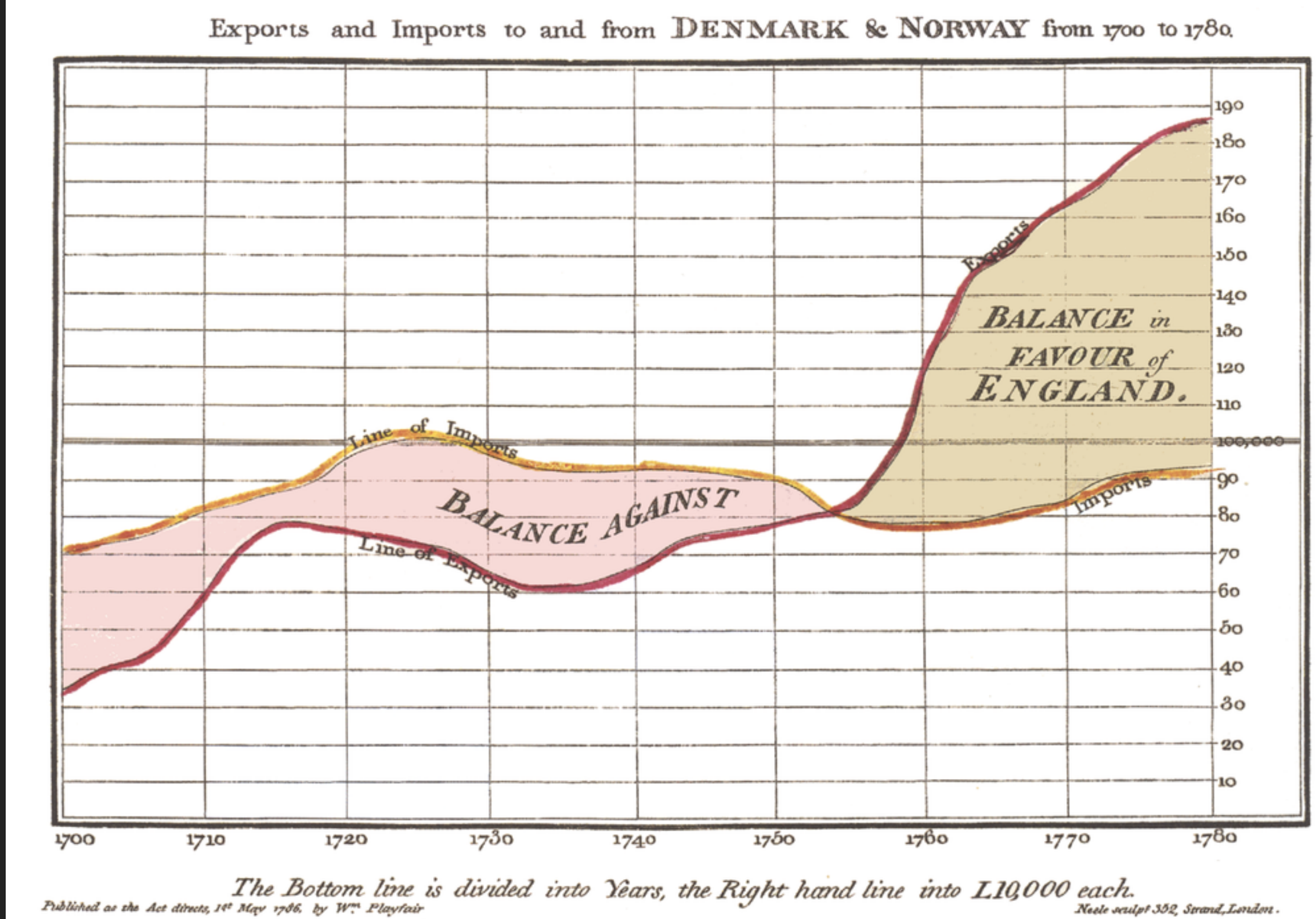
Space and time - fate of Napoleon's army in Russia.
Size of grand army, auxiliary troops, temperature, location, direction.

Dr. John Snow - plotting cholera deaths

In Sept 1854, John Snow plotted cholera deaths (dots) with 11 water pumps (cross). Removed contaminated pump, and epidemic stopped



William Playfair (1759 - 1823)



Another work by William Playfair

IN SUMMARY

GRAPHICAL EXCELLENCE IS THE WELL-DESIGNED PRESENTATION OF INTERESTING DATA – A MATTER OF SUBSTANCE, OF STATISTICS, AND OF DESIGN

Edward Tufte

PRINCIPLES OF GRAPHICAL EXCELLENCE

- Consists of complex ideas communicated with clarity, precision and efficiency
- Is that gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
- always multivariate
- requires telling the truth about the data

"The visual display of Quantitative Information" by Edward Tufte

HOW TO ANALYSE POLITICAL DATA?

Maharashtra results	(288 seats)		
Alliance	**Seats**	**Vote Share(%)**	**Seat Change**
NDA - (BJP+SHS)	161	42.17	-23
UPA - (Cong+NCP)	98	32.53	15
BVA	3	0.67	0
Others	26	23.53	9

Haryana results	(90 seats)		
Alliance	**Seats**	**Vote Share(%)**	**Seat Change**
NDA -(BJP)	40	36.49	-7
UPA -(Cong)	31	28.08	16
INLD	1	2.44	-18
Others	18	32.3	9

Alliances as of Vidhan Sabha elections, 2019

Show results for

Maharashtra

Refine by constituency type clear

Demographics
- SC/ST share above 15%
- Urban
- Rural

By regions
- Khandesh
- Konkan
- Marathwada
- Paschim Maharashtra
- Vidarbha

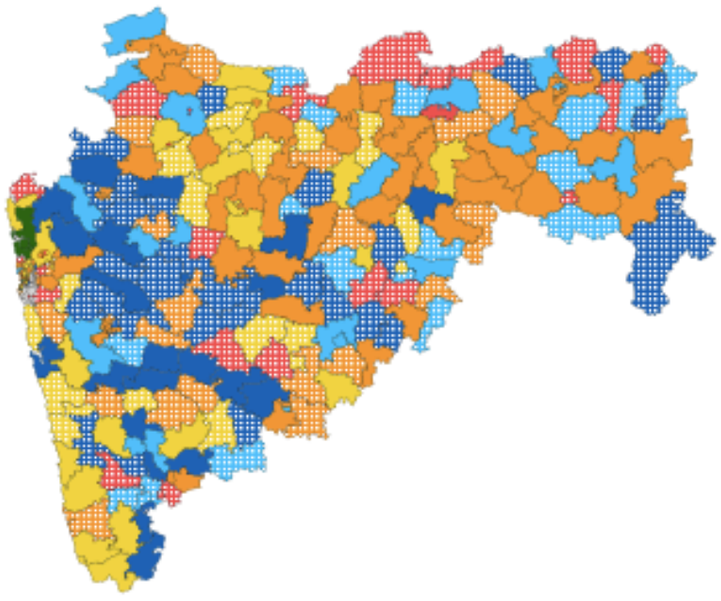
By storylines
- Bhima Koregaon
- Sugar producing

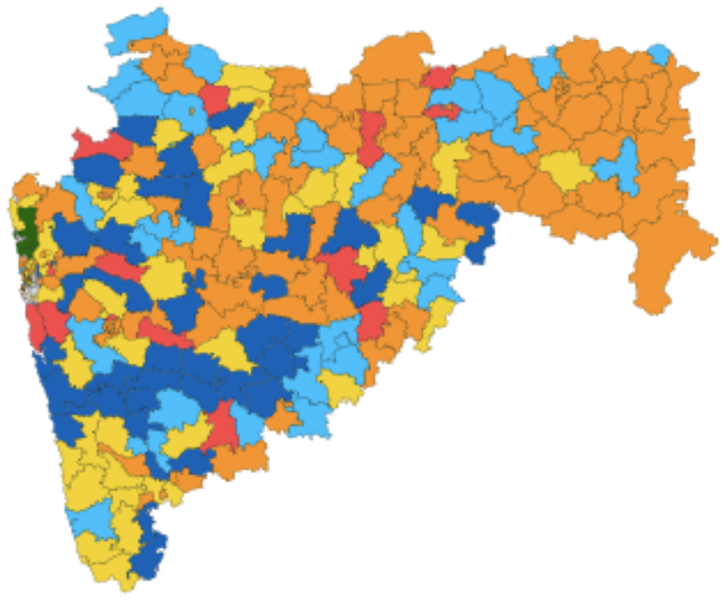
By Census
- Agri dependent
- Owning 2-wheeler
- Owning 4-wheeler
- Access to treated tapwater

Maharashtra | Haryana

Compare 2014 result

2019 2014





− 100 % + RESET

Holds Gains

BJP 105

SHS 56

BJP 122

SHS 63

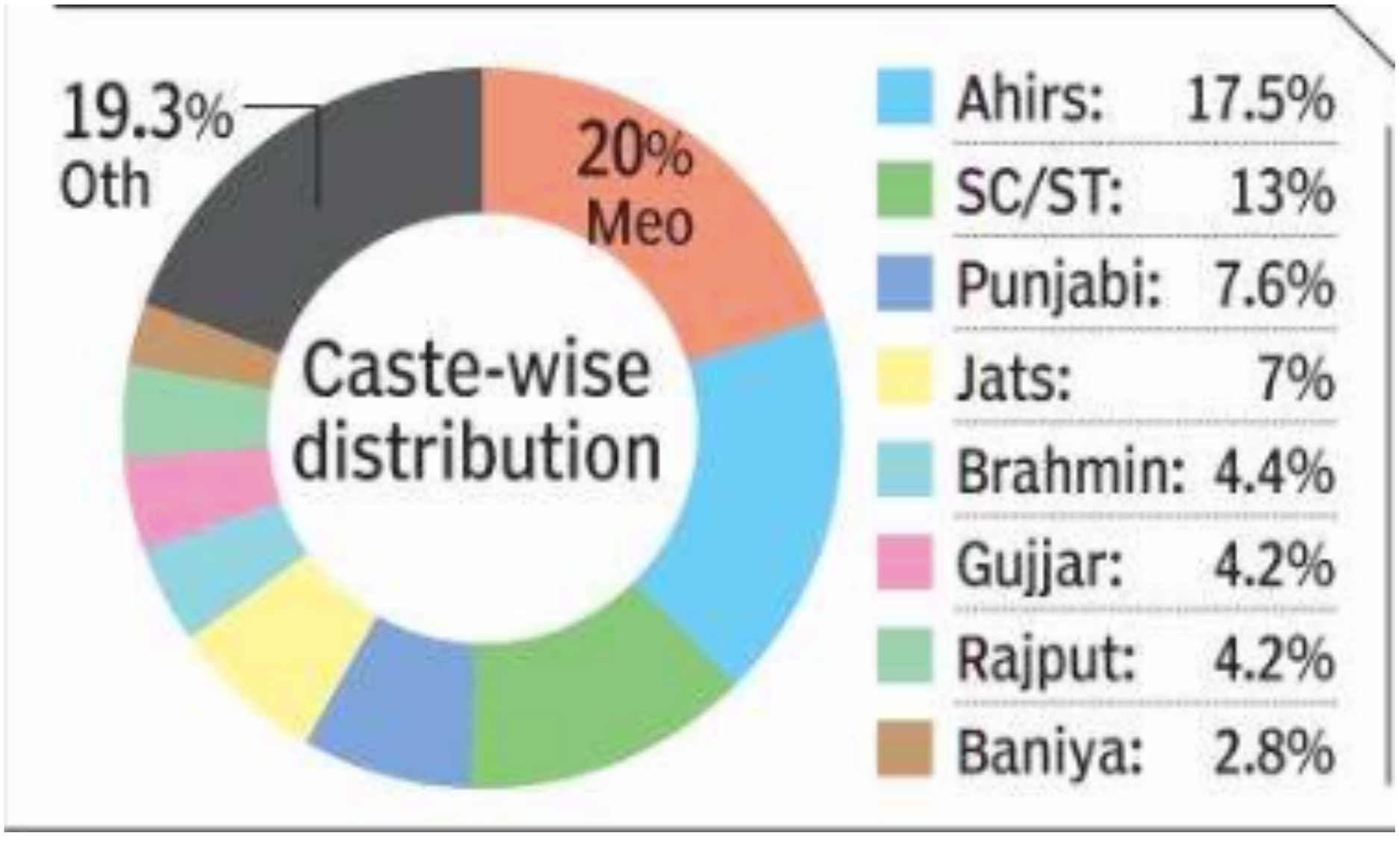
Cong 42

NCP 41

Our latest election coverage - Maharashtra and Haryana elections. Visit HowIndiaLives.com

WHERE'S THE SOURCE?

"Muslims, Yadavs and Scheduled Castes and Tribes, who together account for around 50% of registered voters, will play a key role. Punjabis and Jats are the other major groups in this constituency. As per past trends, almost 80% Yadavs vote for a Yadav candidate.."



https://timesofindia.indiatimes.com/city/gurgaon/caste-religion-city-needs-will-be-key-to-poll-outcome/articleshow/68383176.cms

MULTITUDE OF DATA SOURCES

ELECTION COMMISSION OF INDIA

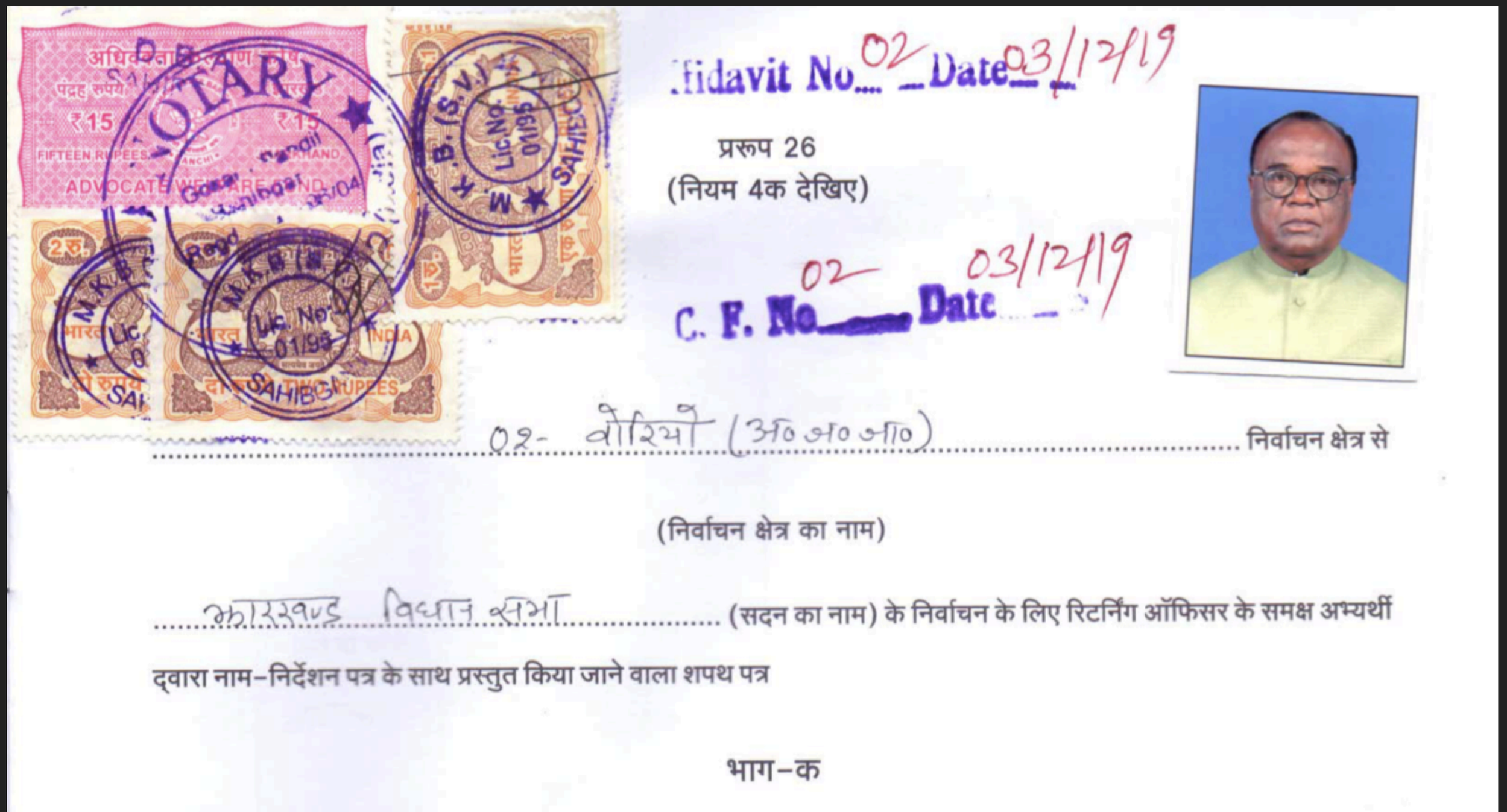
- Counting day
- Past results
- Candidate affidavits
- Form 20 - a powerful data to analyse data
- Voter turnout app
- Political parties funding
- Election expenditure

LOADS OF DATA

BUT THERE ARE ISSUES

POLITICAL DATA

- One data source doesn't talk with another data source.
- Data in scanned PDFs or PDFs, not in machine readable formats.



FEW PRIVATE ORGANISATIONS THAT IS MAKING THIS DATA EASILY AVAILABLE

1. Association for Democratic Rights (ADR) has made affidavits data available in searchable format.
   https://www.myneta.info/_
   Example: Maran of DMK - http://myneta.info/LokSabha2019/candidate.php?candidate_id=6457
2. Trivedi Centre for Political Data - https://tcpd.ashoka.edu.in/

Follow these organisations to track election related data.

DEMOGRAPHIC DATA LAYER TO ELECTION DATA

- Recall - “the greatest number of ideas in the shortest time with the least ink in the smallest space”
- Census is organised as districts, towns, wards and villages. They don’t match with assembly or parliamentary constituencies.
- Delimitation commission matches both Census and Electoral boundaries
- Our story - https://www.livemint.com/elections/assembly-elections/where-the-bjp-fell-behind-and-how-11571939452400.html

OUR EXPERIENCE

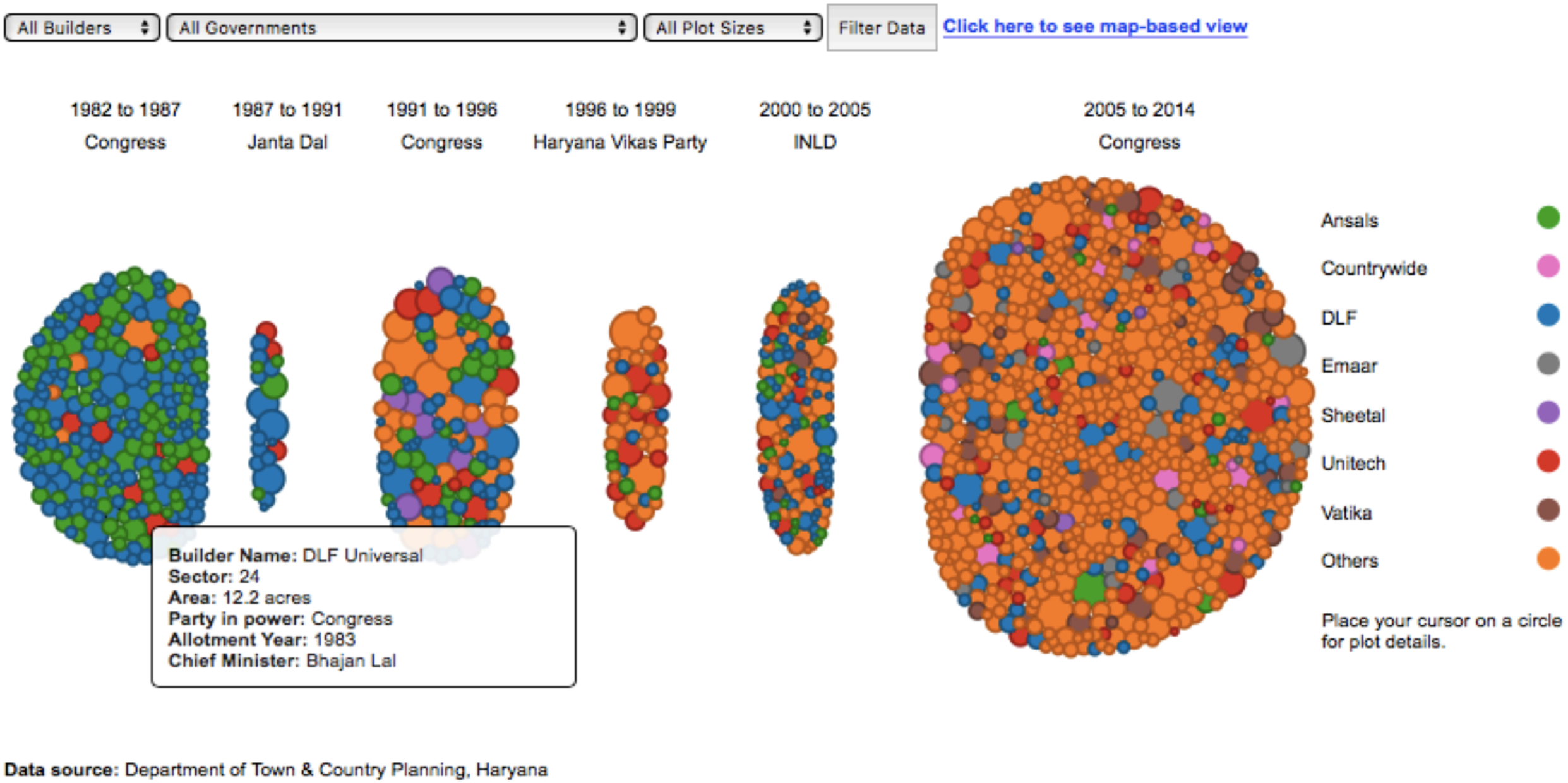
80 + DATASETS, 150 DATA VISUALISATIONS AND
MORE THAN 500 DATA STORIES

ASK ‘SO WHAT’

HOW GURGAON WAS BUILT

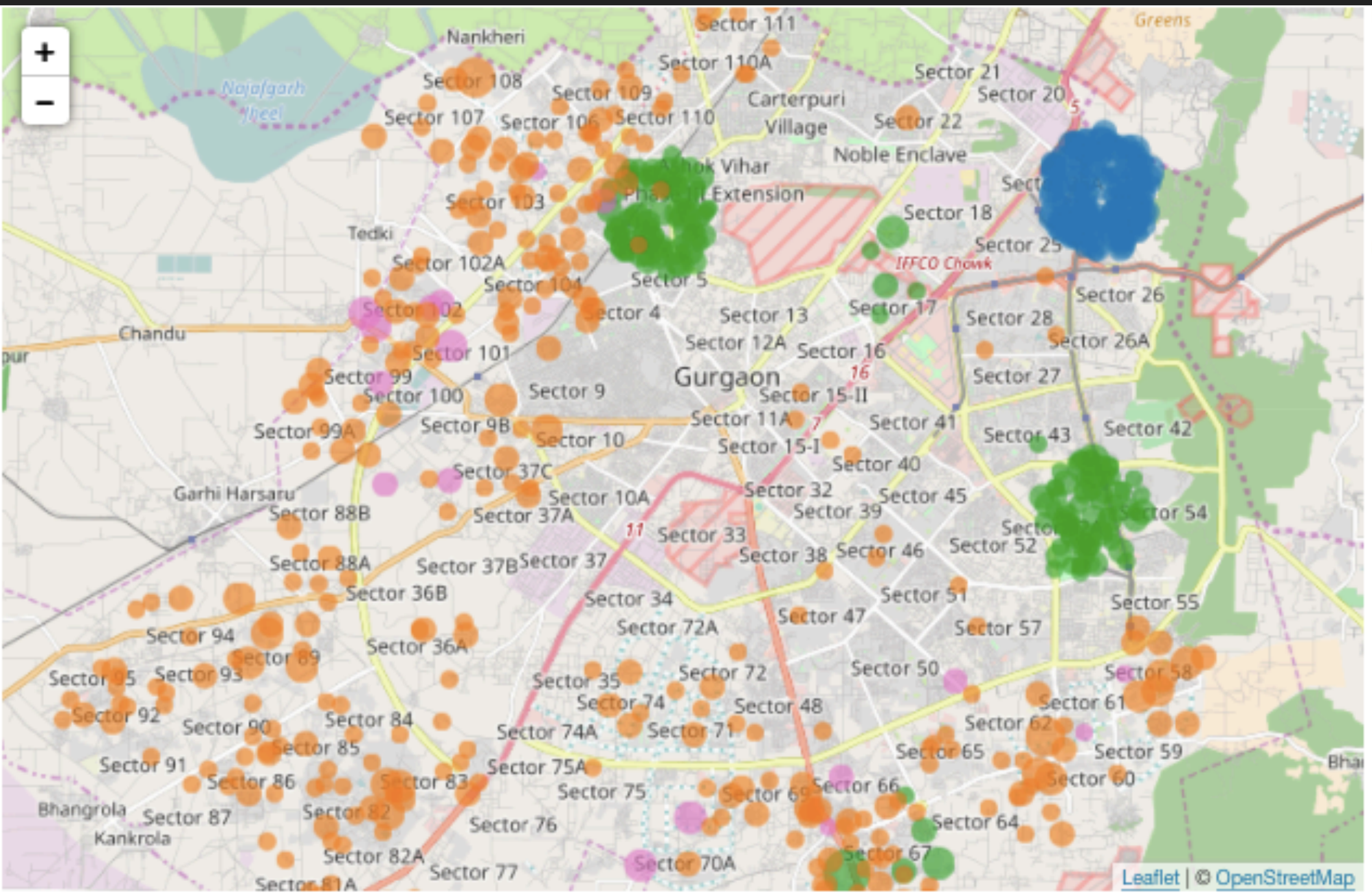
- Released around the same time when a family member of Congress President Sonia Gandhi was accused of corruption
- Data - Town and Country Planning Division, but was in PDF form.
- The choice was to map license allocation with tenure of different ruling parties.

How Gurgaon was built



Data points: developer name, size of land, year of allocation, political parties and when they ruled.

How Gurgaon was built (another version)



Filter by:

All Builders

All Governments

All Plot Sizes

Enter

Slide to see progression:

1981 1992 2014

- Ansals
- Countrywide
- DLF
- Emaar
- Sheetal
- Unitech
- Vatika
- Others

Place your cursor on a circle for plot details.

Click here to go back to bubble graphs

Data source: Department of Town & Country Planning, Haryana

Note: This map assigns a plot to its sector, but not its exact location within the sector. That is because we had the geo-coordinates for the sectors, but not for the plot.

Data points: developer name, size of land, year of allocation, political parties and when they ruled.

IAS OFFICERS AND THE NOTORIOUS TRANSFER CULTURE

- We did this around the time when Ashok Khemka was in the news
- Data - Ministry of Personnel (persmin.nic.in)
- 39,000 postings of 2,600 serving IAS officers who have completed 10 years in service

Haryana is the most transfer-happy state for IAS officers

IAS officers from the Haryana cadre stay, on average, just 11 months in a posting. Their peers at the Centre spend an equally mortifying 22 months. Other states slot into the space in between.
Use this dashboard to see and compare transfer records of serving IAS officers by states, departments, gender and officer.

Search by Cadre

Cadre: All

Department: All

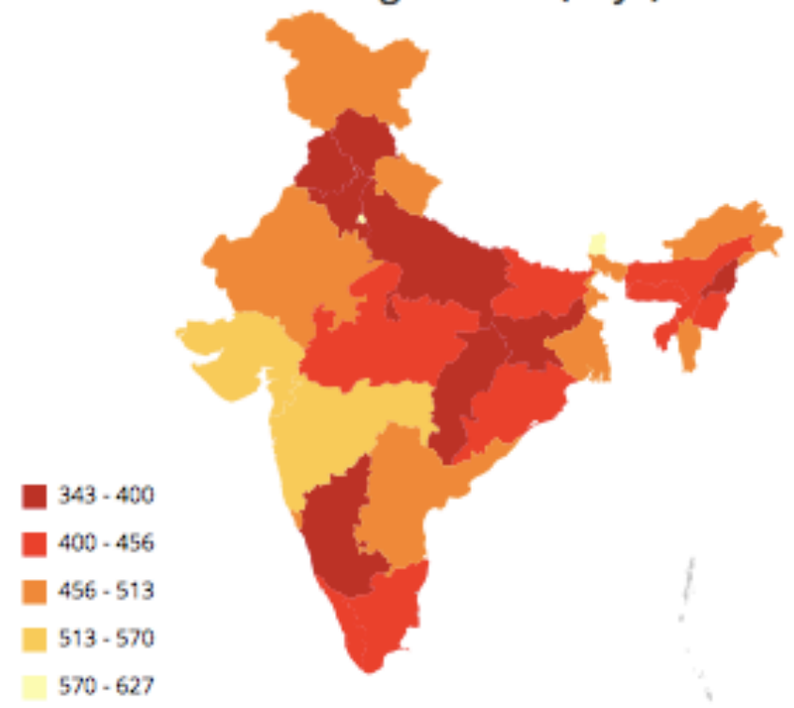
Gender: All

Enter

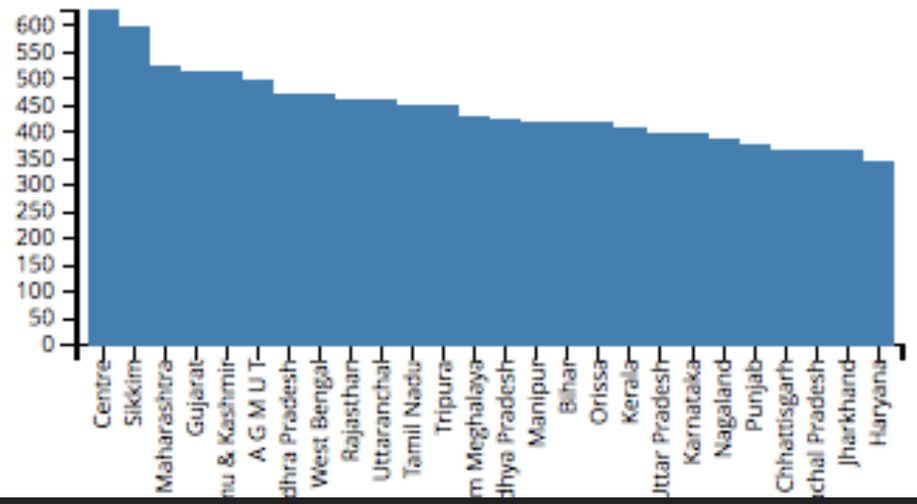
Search by Name

Search by Name

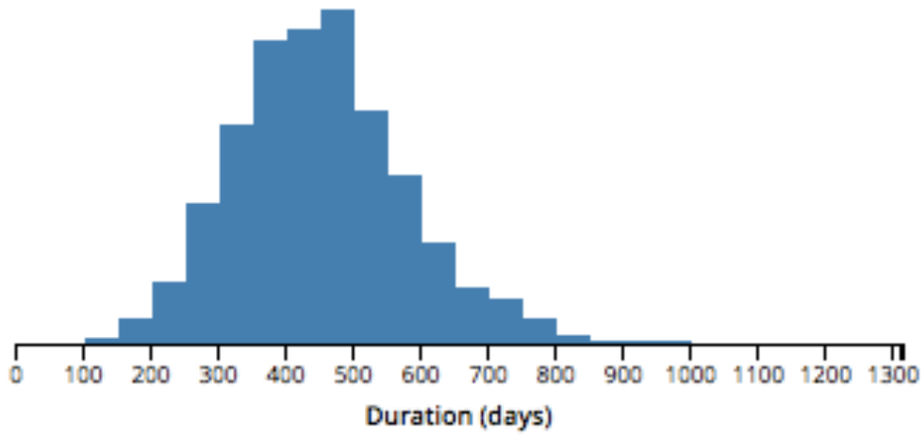
Average tenure (days)



Average tenure (days)



Distribution of officers by avg. tenures



Data points: tenure, officer name, department, location, gender

KEY FINDINGS

- Ashok Khemka is one of the 11 officers to record 40-plus postings. Officers in Haryana, Jharkhand and Himachal average less than a year in a posting. Sikkim, Maharashtra and Gujarat are the best, and even they average barely 18 months per posting
- 10% of postings in 20 key departments last just 33 days to 71 days.
- 22% of IAS officers average less than a year in a posting.
- 68% of IAS officers register 301-550 days per posting.

THE VARIANCE WITH DALITS IN EDUCATION

- The question is to map educational accomplishments of dalits across various states, and compare it with general population
- We wanted to encourage the viewers to compare different pieces of data.
- Data is revealed at different levels of detail

Education levels across dalits

Gender

(•) total ( ) female ( ) male

Location

(•) All ( ) rural ( ) urban

Education Level

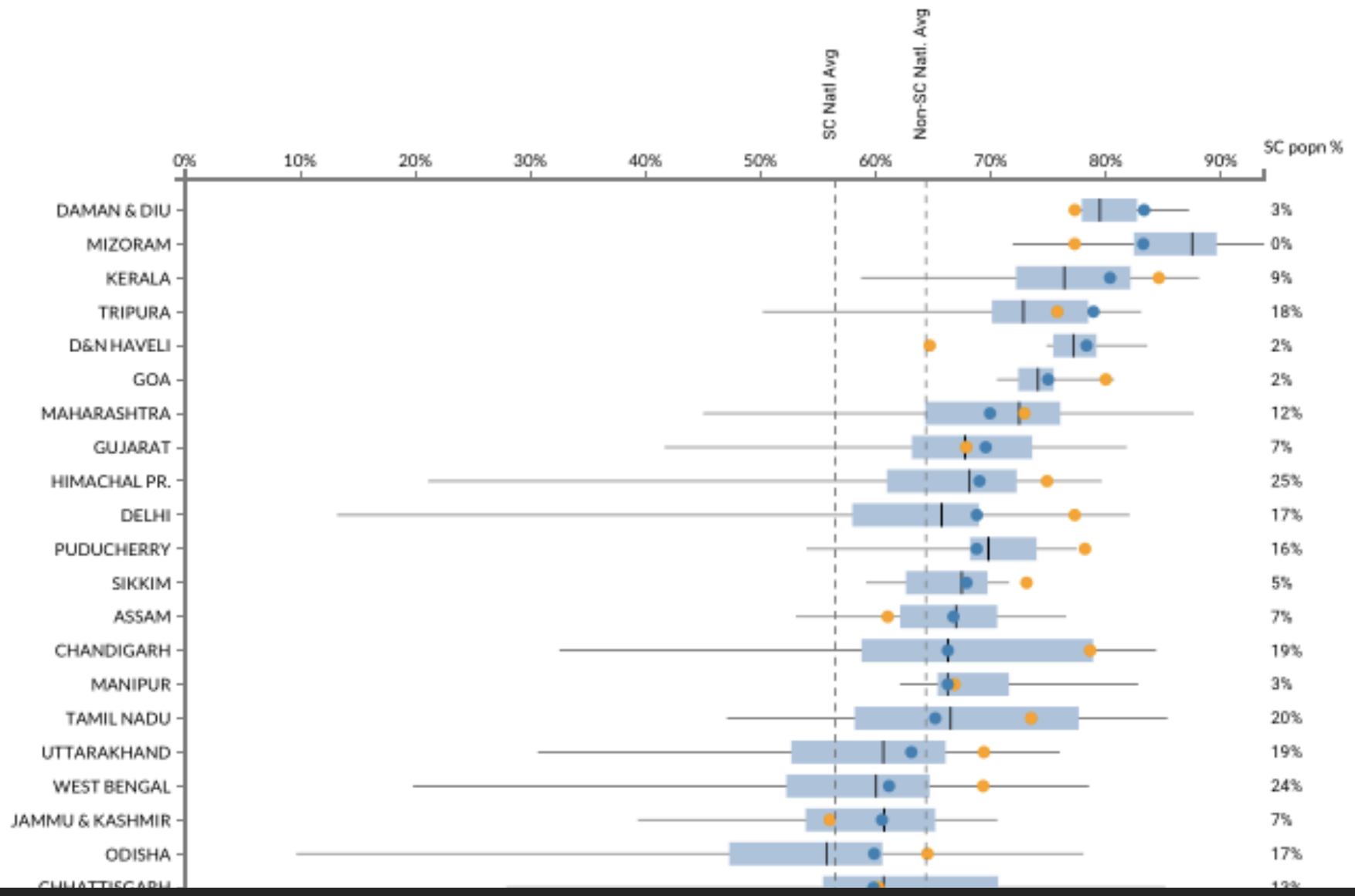
LITERATE ▾

Age Group

ALL AGES ▾

Sort By

(•) Average education level of SC population ( ) Share of SC population ( ) Average education level of non-SC population



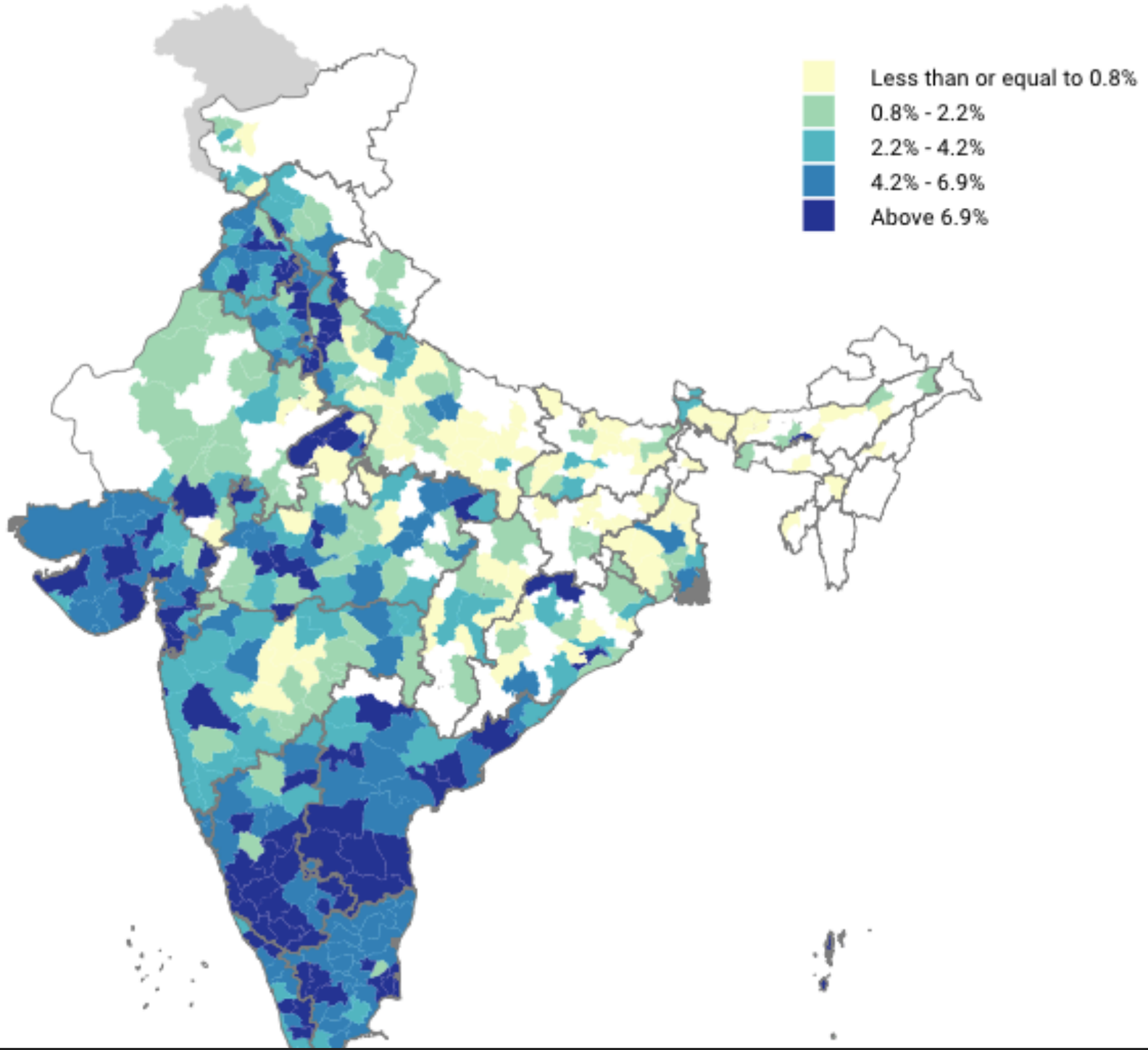
DATA MAPS

LARGE VOLUME OF DATA IN SMALL SPACE

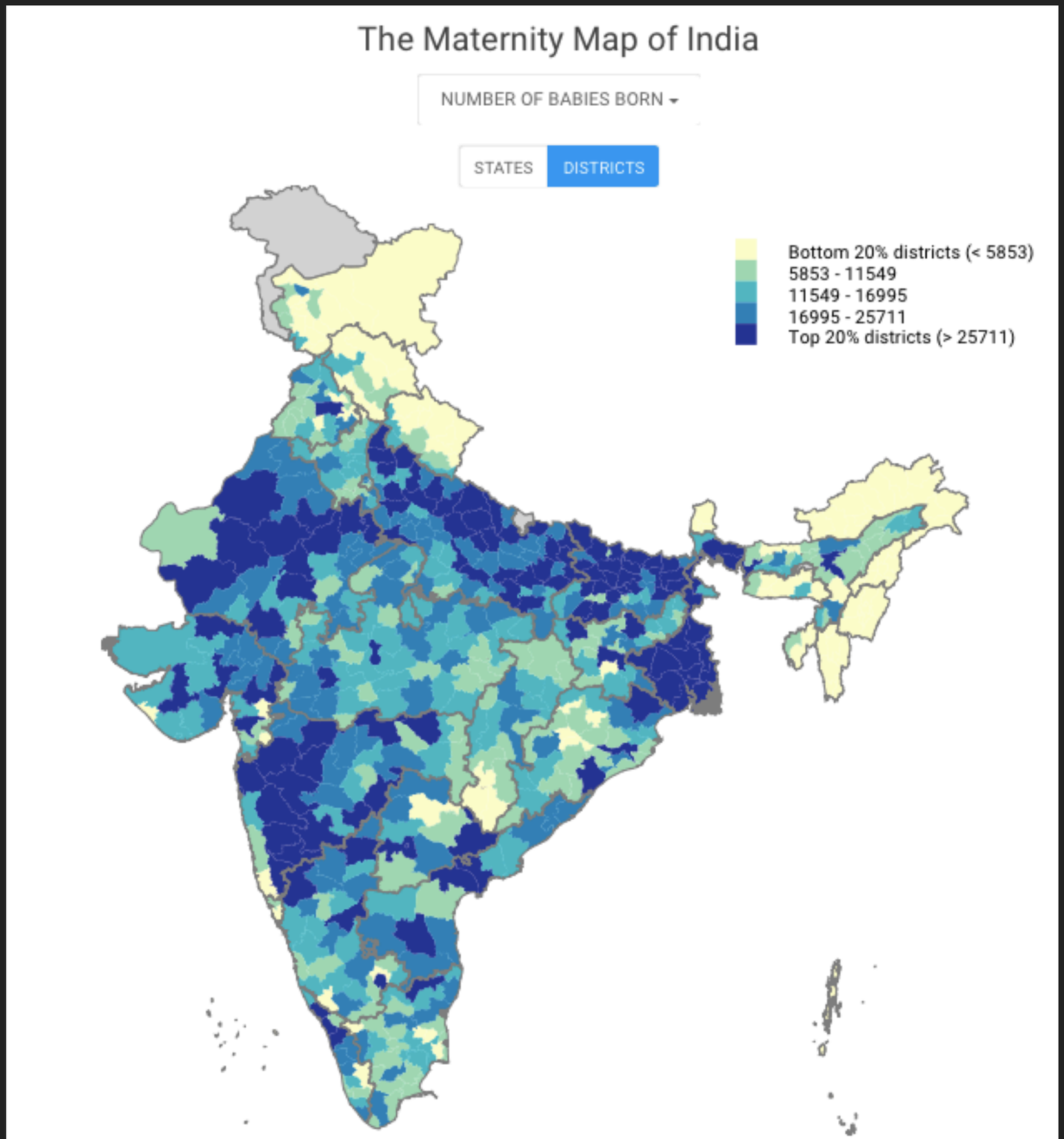
Example 4: Education - mapped 27,000 colleges

IT | ARTS | COMMERCE | EDUCATION | SCIENCE | MANAGEMENT | SOCIAL SCIENCE | OTHERS | ENGINEERING (ALL) | MEDICINE (ALL)

Percent of Students in each district who have chosen stream :Management

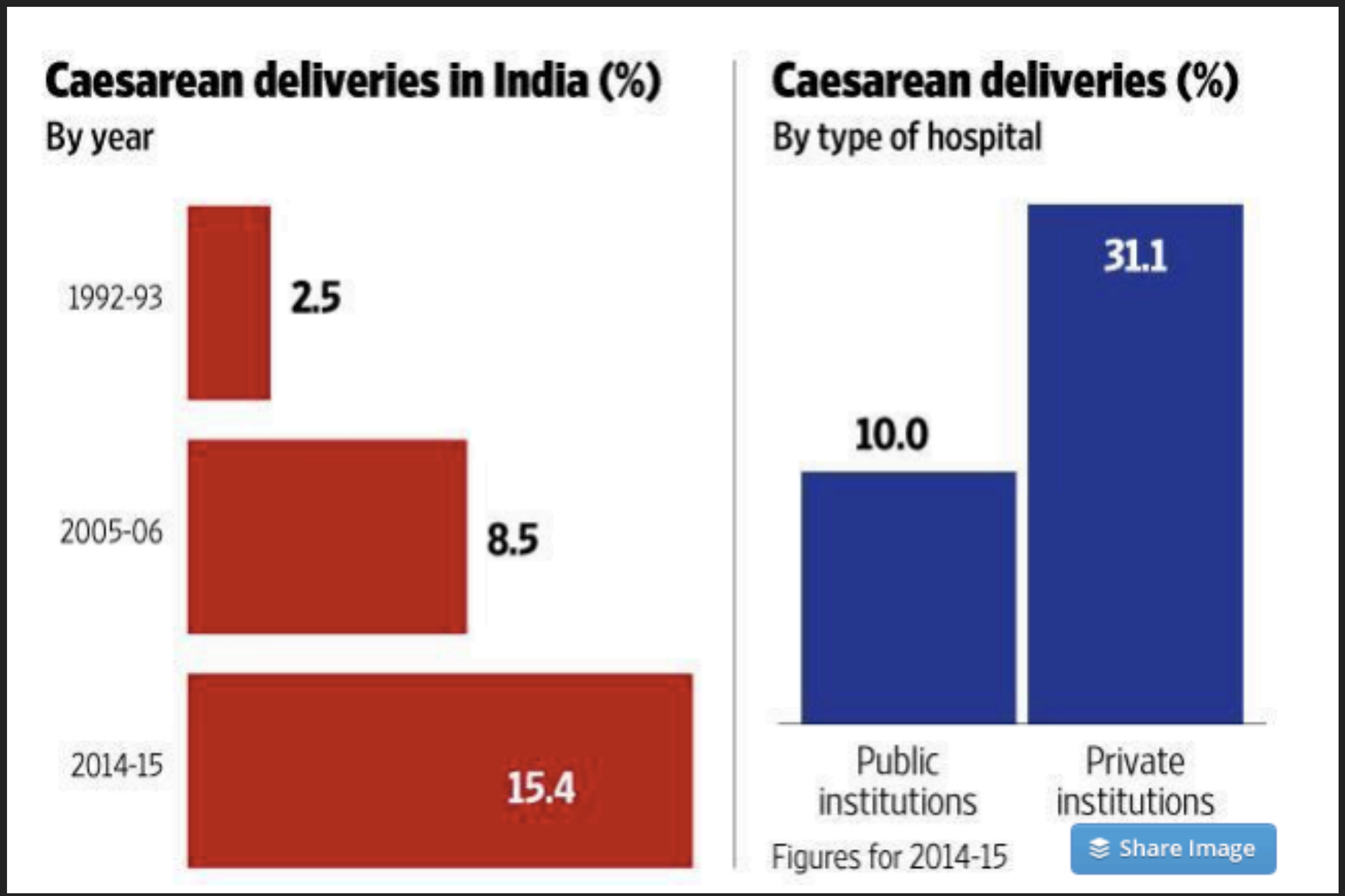


Example 5: Maternity map of India



STATIC GRAPHICS – DOES THE JOB TOO

Alarming rise in caesarean births in India



How infrastructure gains helped gender gains in school

WHERE GIRLS GAINED THE MOST

ELEMENTARY
(Class I to VIII)

GER differential (girls minus boys, percentage points)

	2008-09	2014-15	Swing (percentage points)
Bihar	-16.7	9.9	26.6
Manipur	-7.1	4.9	12.0
Jharkhand	-8.1	3.8	11.9
Puducherry	-4.3	7.6	11.9
Rajasthan	-13.4	-2.7	10.7
Chandigarh	-2.1	7.9	10.1
Arunachal Pradesh	-8.4	0.6	9.1
Uttar Pradesh	1.4	10.0	8.6
Delhi	1.1	8.9	7.8
Meghalaya	1.6	8.7	7.1

WHAT IT TAKES TO DO DATA VIZ

LESSONS FROM HOWINDIALIVES.COM EXPERIENCE

- It takes a digital newsroom to execute a good data viz.
- Four components - data cleaners, coders, journalists/ story tellers and designers
- All have to work together from the start - identification of a data source and how to mine the data.
- 80% of time goes in working with the data. Data viz is grunt work.

TWO WAYS IN APPROACHING DATA

- First approach: Have a question and then use the data to find out the answers
- Second approach: Mine the data and see what it says.

TOOLS WE USE

- ASAP Utilities Excel Add-in (asap-utilities.com)
- Tabula - data from PDF files (tabula.technology)
- DownThemAll - a Firefox plugin to download files from a webpage
- Openrefine - for cleaning data
- Trifacta wrangler - to clean data and basic visualisation
  Additional tips are in the Google Document shared with all.

THANKS

DO FOLLOW US AT @HOWINDIALIVES AND FACEBOOK.COM/HOWINDIALIVES

SUBSCRIBE TO OUR NEWSLETTER WWW.HOWINDIALIVES.COM/NEWS/MAILING-LIST